http://dx.doi.org/10.1090/coll/060/23

CHAPTER 23

# Other combinatorial structures

The ideas of characterizing homomorphism functions, connection ranks, regularity lemmas and limit objects have been extended to several combinatorial structures besides graphs. Some of these extensions are rather involved and deep, like the limit theory of hypergraphs; others can be described as "analogous" (at least after finding the right definitions). Without attempting to be complete, we survey several of these extensions.

## 23.1. Sparse (but not very sparse) graphs

The obvious big gap in our treatment of limits of growing graph sequences is any sequence of graphs with density tending to 0, but maximum degree tending to infinity. Some interesting examples are the point-line incidence graphs of finite projective planes (about $n^{3/2}$ edges, if $n$ is the number of nodes), and $d$-cubes ($n \log n$ edges).

Some work has been done. We have mentioned extensions of the Regularity Lemma to sparser graphs by Kohayakawa [1997], Gerke and Steger [2005], and Scott [2011]). While the case of bounded degree graphs is open, these results are highly nontrivial for sparse (but not very sparse) graphs, and have important applications. They are very likely to play an important role in the limit theory of such graphs. In a substantial paper, Bollobás and Riordan [2009] investigate many of the techniques discussed in this book and elsewhere, mostly from the point of view of extending them from the case of dense graphs to sparser classes.

Lyons [2005] extended the convergence theory of bounded degree graphs to graph sequences with bounded average degree, under a condition called tightness (this guarantees that the sequence of sample distributions has a limit distribution). The following example shows that some condition like this is necessary: the average degree of a subdivision $G'$ of any graph $G$ (dense or not) is bounded by 4. Clearly, to properly describe the limit of the graph sequence $(G'_n)$, the description must contain essentially the same information as the limit of the sequence $(G_n)$. So limits of graphs with bounded average degree are as complex as limits of any graph sequence (dense or sparse).

Graphons and graphings generalize dense graphs and bounded degree graphs, respectively, and they can be considered as the two extremes as far as edge density goes. One common feature is that we can do a random walk on each of them. More precisely, there is a Markov chain on a graphon, as well as on a graphing, and we are going to show that this Markov chain contains all the necessary information about these objects.

Let $W : \Omega^2 \to [0,1]$ be a graphon with density $\omega = t(K_2, W)$. We can define a Markov chain on $\Omega$ by

$$P_u(A) = \frac{1}{d_W(u)} \int\limits_A W(u,v)\,dv$$

(this is defined for almost all $u$). This Markov chain has a stationary distribution, defined by

$$\pi(A) = \frac{1}{\omega} \int\limits_{A\times[0,1]} W(x,y)\,dx\,dy.$$

It is also easy to check that this Markov chain is reversible. The step distribution of this Markov chain is proportional to the integral measure of $W$. Note that the Markov chain does not change if we scale $W$, so we have to remember the "density" $\omega$ if we want to preserve all information about $W$. But the Markov chain together with the density does determine the graphon.

Next, consider a graphing $\mathbf{G}$ on $\Omega$. We can define a Markov chain by

$$P_u(A) = \frac{\deg_A(u)}{\deg(u)}.$$

Then the random walk defined by this chain is just the random walk on this graph in the usual sense. The measure preservation condition (18.2) says that this Markov chain is reversible. A stationary measure of this random walk is $\lambda^*$ (as defined in Section 18.2), its step distribution is $\eta/\omega$. So the step distribution of the Markov chain is the same as the probability measure on the edges of a graphing. The graphing is determined by this Markov chain. It is a fascinating open problem whether Markov chains can be used to define convergence and limit objects for graph sequences that are neither dense nor of bounded degree.

## 23.2. Edge-coloring models

**23.2.1. Edge-connection matrices.** We consider multigraphs with loops. It will be useful to allow a single edge with no endpoints; we call this graph the *circle*, and denote it by $\bigcirc$.

We can define edge-connection matrices that are analogous to the connection matrices defined before: Instead of gluing graphs together along nodes, we glue them together along edges. To be precise, we define a *$k$-broken graph* as a $k$-labeled graph in which the labeled nodes have degree one. (It is best to think of the labeled nodes as not nodes of the graph at all, rather, as points where the $k$ edges sticking out of the rest of the graph are broken off.) We allow that both ends of an edge be broken off.

For two $k$-broken graphs $G_1$ and $G_2$, we define $G_1 * G_2$ by gluing together the corresponding broken ends of $G_1$ and $G_2$. These ends are not nodes of the resulting graph any more, so $G_1 * G_2$ is different from the graph $G_1G_2$ we would obtain by gluing together $G_1$ and $G_2$ as $k$-labeled graphs. We can glue together two copies of an edge with both ends broken off; the result is the circle $\bigcirc$. One very important difference is that while $G_1G_2$ is $k$-labeled, $G_1 * G_2$ has no broken edges any more, and so it is not $k$-broken but 0-broken. This fact leads to considerable difficulties in the treatment of edge models.

For every graph parameter $f$ and integer $k \geq 0$, we define the *edge-connection matrix* $M'(f,k)$ as follows. The rows and columns are indexed by isomorphism types of $k$-broken graphs. The entry in the intersection of the row corresponding to $G_1$ and the column corresponding to $G_2$ is $f(G_1 * G_2)$. Note that for $k = 0$, we have $M(f, 0) = M'(f, 0)$, but for other values of $k$, connection and edge-connection matrices are different.

Let $G$ be a finite graph. An *edge-coloring model* is determined by a mapping $h : \mathbb{N}^q \to \mathbb{R}$, where $q$ is positive integer. We call $h$ the *node evaluation function.* Here we think of $[q]$ as the set of possible edge colors; for any coloring of the edges and $d \in \mathbb{N}^q$, we think of $h(d)$ as the "value" of a node incident with $d_c$ edges with the color $c$ ($c \in [q]$). In statistical physics this is called a *vertex model*: the edges can be in one of several states, which are represented by the color; an edge-coloring represents a state of the system, and (assuming that $h > 0$) $\ln h(d)$ is the contribution of a node (incident with $d_c$ edges with color $c$) to the energy of the state.

There are many interesting and important questions to be investigated in connection with edge-coloring models; we will only consider what in statistical physics terms would be called its "partition function". To be more precise, for an edge-coloring $\varphi : E(G) \to [q]$ and node $v$, let $\deg_c(\varphi, v)$ denote the number of edges $e$ incident with node $v$ with color $\varphi(e) = c$. So the vector $\deg(\varphi, v) \in \mathbb{N}^q$ is the "local view" of node $v$. The *edge-coloring function* of the model is defined by

$$\mathsf{col}(G, h) = \sum_{\varphi:\ E(G)\to[q]} \prod_{v\in V(G)} h\big(\deg(\varphi, v)\big).$$

Recall that we allow the graph $\bigcirc$ consisting of a single edge with no endpoints; by definition, $\mathsf{col}(\bigcirc, h) = q$. We also allow that $q = 0$, in which case $\mathsf{col}(G, h) = 1$ if $G$ has no edges, and $\mathsf{col}(G, h) = 0$ otherwise. We could of course allow complex valued node evaluation functions, in which case the value of the edge-coloring function can be complex.

EXAMPLE 23.1 **(Number of perfect matchings)**. The number of perfect matchings can be defined by coloring the edges by two colors, say black and white, and requiring that the number of black edges incident with a given node be exactly one. This means that this number is $\mathsf{col}(., h)$, where $h : \mathbb{N}^2 \to \mathbb{R}$ is defined by $h(d_1, d_2) = \mathbb{1}(d_1 = 1)$. The number of all matchings could be expressed similarly. ◆

EXAMPLE 23.2 **(Number of 3-edge-colorings)**. This number is $\mathsf{col}(., h)$, where $h : \mathbb{N}^3 \to \mathbb{R}$ is defined by $h(d_1, d_2, d_3) = \mathbb{1}(d_1, d_2, d_3 \le 1)$. ◆

EXAMPLE 23.3 **(Spectral decomposition of a graphon)**. Recall the definition (7.18) and expression (7.25) for $t(F, W)$ in terms of the spectrum of $T_W$. We can consider $\chi$ as a coloring of $E(F)$ with colors $1, 2, \dots$. Then $M_\chi(v)$ depends only on the numbers of edges with different colors, and so we can write $M_\chi(v) = h\big(\deg(\chi, v)\big)$, and we get

$$t(F, W) = \mathsf{col}(G, \lambda, h) = \sum_{\chi:\ E(G)\to[q]} \prod_{e\in E(F)} \lambda_{\chi(e)} \prod_{v\in V(G)} h\big(\deg(\chi, v)\big).$$

However, this is not a proper edge-coloring model, since the value of the circle, which is the number of colors, is infinite in general. ◆

The following facts about the edge-connection matrices of edge-coloring functions are easy to prove along the same lines as Proposition 5.64:

PROPOSITION 23.4. *For every edge-coloring model $h : \mathbb{N}^q \to \mathbb{R}$, the graph parameter $\mathsf{col}(., h)$ is multiplicative, its edge-connection matrices $M'(f, k)$ are positive semidefinite, and* $\mathrm{rk}(M'(f, k)) \le q^k$. □

B. Szegedy [2007] showed that the first two of these properties suffice to give a characterization of edge-coloring functions.

THEOREM 23.5. *A graph parameter $f$ can be represented as an edge-coloring function if and only if it is multiplicative and $M'(f, k)$ is positive semidefinite for all $k \ge 0$.* □

The proof of this theorem is quite involved and not reproduced here; it is based on ideas similar to those used in Section 6.6 to prove Theorem 5.57, but using quite a bit more involved tools: The use of the Nullstellensatz and simple semidefiniteness arguments must be replaced by real versions of the Nullstellensatz (Positivstellensatz), and the simple symmetry arguments must be replaced by deeper results from the representation theory of algebras. (As a historical comment, the proof of Theorem 23.5 came first, and Schrijver's proof of Theorem 5.57 was motivated by this.)

Draisma, Gijswijt, Lovász, Regts and Schrijver [2012] give characterizations of complex valued edge-coloring functions. Let us state without proof a result of Schrijver [2012], which shows that a condition on the growth of the edge-connection rank (along with minor other constraints) can characterize complex edge-coloring models.

THEOREM 23.6. *A complex valued graph parameter $f$ is an edge-coloring function of a complex model if and only if it is multiplicative, $f(\bigcirc)$ is real, and* $\mathrm{rk}\big(M'(f, k)\big) \le f(\bigcirc)^k$ *for every $k$.* □

**23.2.2. Tensor algebras.** There is a surprisingly close connection between edge-coloring models and rather general multilinear algebra. Given an edge-coloring model (with color set $[q]$), we can think of the nodes as little gadgets with wires (or legs) sticking out corresponding to the edges incident with it. If we assign colors to the wires, the gadget outputs a number (this could be real or complex). The graph parameter defined by this edge-coloring model is the expectation of the product of these numbers, one for each node, where the edge-coloring is random.

Formulating the question like this, we see two restrictions that look artificial:

— There is only one gadget for each degree. Why not have several?

— We have assumed that the output of a gadget depends only on the number of legs with each color; in other words, it is invariant under the permutation of the legs. Why not drop this condition?

If we relax these conditions, then every gadget would be described by a real array $(H_{i_1,\dots,i_d} : i_r \in [q])$, where $d$ is the number of legs. In other words, the gadget is described by a tensor with $d$ slots over $\mathbb{R}^q$. Furthermore, we have to indicate for each node $v$ which gadget is sitting there, and how its legs correspond to the edges incident with $v$.

In more mathematical terms, we have a graph where a tensor is associated with every node, and an index associated with every edge, so that the slots (indices) of the tensor correspond to the edges incident with the node. Let us call such a graph a *tensor network*. The corresponding graph parameter is evaluated by taking the product of these tensors, and then summing over all choices of the indices. Note that every index occurs twice, so we could call this "tracing out" every index.

These tensor networks play an important role in several areas of physics, but we can't go into this topic in this book.

This setup allows for a more general construction. If we have a tensor network with $k$ broken edges, then the value associated with the graph will depend on the color of these edges, in other words, it will be described by an array $(A_{i_1,\dots,i_k} : i_r \in [q])$. So the graph with $k$ broken edges can be considered as a gadget itself.

We can break down the procedure of assembling a tensor network from the gadgets (with or without broken edges) into two very simple steps:

(a) We can take the disjoint union of two gadgets; if the gadgets have $k$ and $l$ legs, respectively, the union has $k+l$ legs. In terms of multilinear algebra, this means to form the tensor product of two tensors.

(b) We can fuse two legs of a gadget. If $(A_{i_1,\dots,i_k} : i_r \in [q])$ is the tensor describing the gadget, and (say) we fuse legs $k-1$ and $k$, then we get the tensor

$$B_{i_1,\dots,i_{k-2}} = \sum_{j\in[q]} A_{i_1,\dots,i_{k-2},j,j}.$$

In multilinear algebra slang, we trace out the last two indices.

It is easy to see that with these operations, we can construct every tensor network with or without broken edges, and we get the corresponding tensor.

Supposing that we have a starting kit of gadgets, we can look at the set of all tensors that can be realized by assembling tensor graphs with broken edges from these gadgets. In the spirit of linear algebra, we take all linear combinations of the obtained tensors with the same number of slots. Every tensor obtained this way will be called an *assembled tensor*.

It is clear from (a) and (b) above that the set of assembled tensors has the following structure: For every $k$, there is a linear space $\mathcal{T}_k$ of tensors over $\mathbb{R}^q$ with $k$ slots. For every $A \in \mathcal{T}_k$ and $B \in \mathcal{T}_l$, the tensor product $A \otimes B \in \mathcal{T}_{k+l}$. For every $A \in \mathcal{T}_k$, and any two indices in $A$, tracing out these two indices results in a tensor in $\mathcal{T}_{k-2}$. We call such a set of tensors a *traced tensor algebra*.

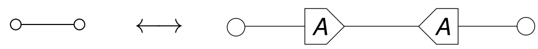
Conversely, every traced tensor algebra arises as the set of assembled tensors: for every number $k$ of slots, we select a basis of the space $\mathcal{T}_k$, and use the resulting set of tensors as the starting kit.

It is quite fruitful to use this connection; one can obtain results that are new both for graphs and for tensor algebras. We describe one important result with combinatorial connections.

Given a starting kit $K$, how can we decide about a tensor whether it can be assembled from this kit? In other words, is it contained in the traced tensor algebra generated by $K$? A beautiful answer to this question was found by Schrijver [2008a], which we describe in graph-theoretic terms (the proof uses the representation theory of algebras, and we do not give it here; cf. also Schrijver [2008b, 2009]).

Recall that we work over a fixed vector space $\mathbb{R}^q$. Every $q\times q$ real matrix $A$ is a gadget in itself, with two legs. If it is symmetric, then the legs are interchangeable, but in general we have to talk about a "left leg" (corresponding to the row index) and a "right leg" (corresponding to the column index). Connecting the gadgets for matrices $A$ and $B$ in series gives a gadget representing the matrix $AB$.

Orthogonal matrices will play a special role. One observation is that if $A$ is orthogonal, then $AA^{\mathsf{T}} = I$ (the identity matrix), and so if we have a gadget graph with no broken edges, and replace any edge by a path of length 3 with $A$ and $A^{\mathsf{T}}$ sitting on it:




If we replace every edge by this path of length 3, then the value of the graph does not change. However, we can group together every original gadget $B$ with the orthogonal matrices next to, to get a gadget $B^A$, which—in multilinear algebra terms—is obtained from $B$ by applying the linear transformation $A$ to every slot. If we replace every gadget $B$ in the kit by $B^A$, then the value of the tensor network does not change.

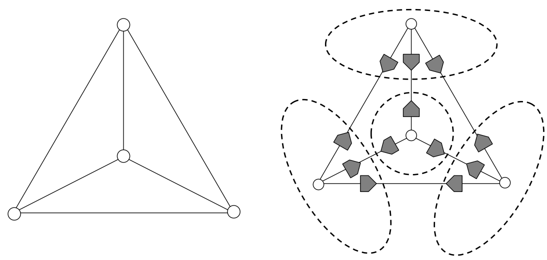


FIGURE 23.1. Replacing every edge by a path with the same orthogonal transformation at both inner nodes (just facing the opposite direction), and regrouping does not change the value.

Now consider a tensor network with broken edges. If we replace every tensor $B$ in the kit by $B^A$, then the matrices $A$ and $A^{\mathsf{T}}$ along the unbroken edges still cancel each other, but on the broken edges, one copy still remains. In other words, if we apply the same orthogonal transformation to every slot of every tensor in the kit, then the tensor defined by a tensor network with broken edges undergoes the same transformation.

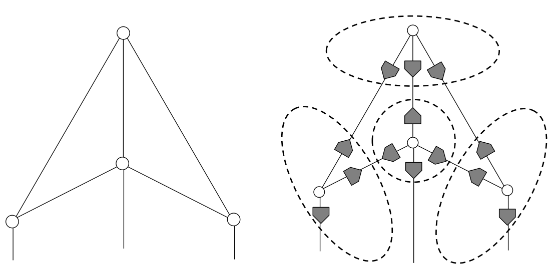


FIGURE 23.2. Applying the same orthogonal transformation to all slots of all tensors in the kit results in applying the same orthogonal transformation to the slots of the assembled tensor.

In particular, if all tensors in the kit have the property that a particular orthogonal transformation applied to all their slots leaves them invariant, then the same holds for every assembled tensor. The theorem of Schrijver [2008a] asserts that this is the only obstruction to assembling a given tensor.

THEOREM 23.7. *Let $\mathcal{T}$ be a traced tensor algebra generated by a set $\mathcal{S}$ of tensors, including the identity tensor $\mathbb{1}(i=j)$ $(i,j \in [q])$. Then a tensor $T$ is in $\mathcal{T}$ if and only if it is invariant under every orthogonal transformation that leaves every tensor in $\mathcal{S}$ invariant.* □

The special case when the generating tensors are symmetric describes edge-coloring models. This can be viewed as an analogue of Theorem 6.38, with the role of the edges and nodes interchanged. Regts [2012] showed how Theorem 23.7 yields an exact formula for the edge-connection rank of edge-coloring models.

EXAMPLE 23.8 **(Number of perfect matchings revisited)**. The tensor model for this graph parameter is a bit more complicated than in Example 23.1. We have 2 edge colors (which will be convenient to call 0 and 1), so we work over $\mathbb{R}^2$; but we need to specify a tensor for every degree $d$, expressing that exactly one edge is black:

$$T_{i_1,\dots,i_d} = \mathbb{1}(i_1 + \dots + i_d = 1).$$

It is easy to see that no orthogonal transformation, applied to all slots, leaves this tensor invariant, so it follows from Theorem 23.7 that every tensor can be assembled from this kit. (We note that the tensor is invariant under permuting the slots; however, this symmetry is not preserved under composition of tensor networks.) ♦

EXAMPLE 23.9 **(Number of 3-edge-colorings revisited)**. To construct a tensor model for the number of 3-edge-colorings, we work over $\mathbb{R}^3$. We again need to specify a tensor for every degree expressing that the edges have different colors:

$$T_{i_1,\dots,i_d} = \mathbb{1}(i_1,\dots,i_d \text{ are different})$$

(for $d > 3$, we get the 0 tensor). Permuting the colors (i.e., the coordinates in the underlying vector space $\mathbb{R}^3$) leaves this tensor invariant, and these are the only orthogonal transformations of $\mathbb{R}^3$ with this property. Theorem 23.7 implies that a tensor is invariant under the permutations of the coordinates of $\mathbb{R}^3$ if and only if it can be assembled from this kit. ♦

## 23.3. Hypergraphs

When talking about generalizing results on graphs, the first class of structures that comes to mind is hypergraphs (at least to a combinatorialist). So it is perhaps surprising that to extend the main concepts and methods developed in this book (quasirandomness, limit objects, Regularity Lemma, and Counting Lemma) to hypergraphs is highly nontrivial. Even the "right" formulation of the Regularity Lemma took a long time to find, and in the end both the Regularity Lemma and the limit object turned out quite different from what one would expect as a naive generalization. Nevertheless, the issue is essentially solved now, thanks to the work of Chung, Elek, Graham, Gowers, Rödl, Schacht, Skokan, Szegedy, Tao and others. A full account of this work would go way beyond the possibilities of this book, but we will give a glimpse of the results.

By an *$r$-uniform hypergraph*, or briefly *$r$-graph*, we mean a pair $H = (V, E)$, where $V = V(H)$ is a finite set and $E = E(H) \subseteq \binom{V}{r}$ is a collection of $r$-element subsets. The elements of $V$ are called *nodes*, the elements of $E$ are called *edges*. So 2-graphs are equivalent to simple graphs. We can define the homomorphism number $\mathsf{hom}(G,H)$ of an $r$-graph $G$ into an $r$-graph $H$ in the natural way, as the number of maps $\varphi: V(G) \to V(H)$ for which $\varphi(A) \in E(H)$ for every $A \in E(G)$. The *homomorphism density* of $G$ in $H$ is defined as one expects, by the formula

$$t(G,H) = \frac{\mathsf{hom}(G,H)}{|V(G)|^{|V(H)|}}$$

Quasirandomness can be defined by generalizing the condition on the density of quadrilaterals. We need to define a couple of special hypergraph classes. Let $K_n^r$ denote the complete $r$-uniform hypergraph on $[n]$ (i.e., $E(K_n^r) = \binom{[n]}{r}$). Let $L_k^r$ be the "complete $r$-partite hypergraph" defined on the node set $V_1 \cup \dots \cup V_r$, where the $V_i$ are disjoint $k$-sets, and the edges are all $r$-sets containing exactly one element from each $V_i$. Clearly $t(K_r^r, H) = t(L_1^r, H)$ is the edge density of $H$. It is not hard to prove that $t(L_k^r, H) \ge t(K_r^r, H)^{k^r}$ for every $H$ (this generalizes inequality 2.9 from the Introduction). We define the *quasirandomness* of $H$ as the difference $\mathsf{qr}(H) = t(L_2^r, H) - t(K_r^r, H)^{2^r}$.

A sequence $(H_n)$ of hypergraphs is called *quasirandom* with density $p$ if $t(K_r^r, H_n) \to p$ and $\mathsf{qr}(H_n) \to 0$, or equivalently, $t(L_2^r, H_n) \to p^{2^r}$. It was proved by Chung and Graham [1989] that this implies that $t(G, H_n) \to p^{e(G)}$ for every $r$-graph $G$, so the equivalence of conditions (QR2) and (QR3) for quasirandomness in the Introduction (Section 1.4.2) generalizes nicely.

As a first warning that not everything extends in a straightforward way, let us try to generalize (QR5). A first guess would be to consider disjoint sets $X_1, \dots, X_r \subseteq V$, and then stipulate that the number of edges with one endpoint in each of them is $p|X_1| \dots |X_r| + o(n^r)$. (For simplicity of presentation, we assume that $v(H_n) = n$.) This property is indeed valid for every quasirandom sequence, but it is strictly weaker than quasirandomness. It is not well-defined what the "right" generalization is; we state one below, which is a version of a generalization found by Gowers. Several other equivalent conditions are given by Kohayakawa, Rödl and Skokan [2002].

PROPOSITION 23.10. *A sequence $(H_n)$ of hypergraphs is quasirandom with density p if and only if for every $(r-1)$-graph $G_n$ on $V(H_n)$, the number of edges of $H_n$ that induce a complete subhypergraph in $G_n$ is $t(K_r^{r-1}, G_n)t(K_r^r, H_n)\binom{n}{r} + o(n^r)$.* □

In the case of simple graphs ($r = 2$), let $H_n$ be a simple graph with edge density $p$. The 1-graph $G_n$ means simply a subset of $V(H_n)$, and $K_2^1$ is just a 2-element set. So the condition says that the number of edges of the graph $H_n$ induced by the set $G_n$ is asymptotically

$$t(K_2^1, G_n)t(K_2^2, H_n)\binom{n}{2} = \Big(\frac{|G_n|}{n}\Big)^2 \frac{2\mathsf{e}(H_n)}{n^2}\binom{n}{2} \sim p\binom{|G_n|}{2},$$

and so we get condition (Q4). For general $r$, the condition can be rephrased as follows: for a random $r$-set $X \subseteq V$, the events that $X$ is complete in $G_n$ and $X$ is an edge in $H_n$ are asymptotically independent.

The last remark takes us to another complication.

EXAMPLE 23.11. Let $\mathbb{G}(n, 1/2)$ be a random graph and let $\mathbb{T}_n$ denote the 3-graph formed by the triangles in $\mathbb{G}(n, 1/2)$. Then $\mathbb{T}_n$ is a 3-graph with density 1/8, which is random in some sense, but it is very different from the random 3-graph $\mathbb{H}_n$

on $[n]$ obtained by selecting every edge independently with probability 1/8. In fact, the sequence $(\mathbb{H}_n)$ is quasirandom with probability 1 (this is not hard to see), while $\mathbb{T}_n$ has a very small intersection with every quasirandom 3-graph by Proposition 23.10. Also, $\mathbb{T}_n$ has some special features, like no 4-set of nodes contains exactly 3 edges of $H_n$.

On the other hand, $\mathbb{T}_n$ is totally homogeneous. It has no special global structure; more concretely: on any two disjoint $k$-sets we see independent copies of the same random hypergraph. If we want to generalize the Regularity Lemma, it has to reflect the difference between $\mathbb{T}_n$ and $\mathbb{H}_n$, and similarly for the generalization of the notion of graphons. Which of these sequences should tend to a constant function? ♦

We show how to overcome this difficulty, starting with the construction of the limit object. We say that a sequence of $r$-graphs $(H_n)$ is *convergent*, if $v(H_n) \to \infty$ and $t(F, H_n)$ has a limit as $n \to \infty$ for every $r$-graph $F$. Let $t(F)$ denote this limit. How to represent this limit function, in other words, what is the hypergraph analogue of a graphon? The natural guess would be a symmetric $r$-variable function $W : [0,1]^r$, which would represent the limit by

$$t(F, W) = \int_{[0,1]^r} \prod_{\{i_1,\dots,i_r\}\in E(F)} W(x_{i_1}, \dots, x_{i_r^2})\, dx.$$

The example of the hypergraphs $\mathbb{H}_n$ and $\mathbb{T}_n$ above show that this cannot be right. The only reasonable candidate for their limit object would be the function $W \equiv 1/8$, which represents correctly the limiting densities for the sequence $\mathbb{H}_n$, but not for the sequence $\mathbb{T}_n$. We could make life even more complicated, and consider the intersection $\mathbb{H}_n \cap \mathbb{T}_n$, which is a random 3-graph with expected density 1/64, and the limiting densities are even more complicated.

For $r > 3$, one could construct a whole zoo of homogeneous random hypergraphs, generalizing the construction of $\mathbb{H}_n$ and $\mathbb{T}_n$. After several steps of generalization, one arrives at the following: we generate a random coloring of $K_n^j$ for every $0 \le j \le r$ (with any number of colors). To decide whether an $r$-subset $X \subseteq [n]$ should be an edge, we look at the colors of its subsets, and see if this coloring belongs to some prescribed family of colorings of $2^X$. (We assume that the prescribed family is invariant under permutations of $X$.)

While this example warns us of complications, it also suggests a way out: we describe the limit not in the $r$-dimensional but in the $2^r$-dimensional space. In fact, the limit object turns out to be a subset, rather than a function, which is a gain (it is of course very little relative to the increase in the number of coordinates).

Consider the set $[0,1]^{2^{[r]}}$ (so we have a coordinate $x_I$ for every $I \subseteq [r]$; the coordinate for $\emptyset$ will play no role, we can think of it as 0). Let us note that the symmetric group $S_r$ acts on the power set $2^{[r]}$, and hence also on $[0,1]^{2^{[r]}}$. Let $U \subseteq [0,1]^{2^{[r]}}$ be a measurable set that is invariant under the action of $S_r$. We call such a set a *hypergraphon*.

For every hypergraphon $U$, we define the density of an $r$-graph $F$ as follows. We assign independent random variables $X_S$, uniform in $[0,1]$, to every subset $S \subseteq V(F)$ with $|S| \le r$. For every edge $A = \{a_1, \dots, a_r\} \in E(F)$, and every $I \subseteq [r]$, we denote by $A_I$ the subset $\{a_i : i \in I\}$, and we consider the point $X(A) \in [0,1]^{2^{[r]}}$ defined by $(X(A))_I = X_{A_I}$ (this depends on the ordering of $A$, but this will not matter thanks to our symmetry assumption about $U$). Now we define

$$t(F,U) = \mathsf{P}\big(X(A) \in U \text{ for all } A \in E(F)\big).$$

To illuminate the meaning of this formula a little, consider the case $r = 2$. Then we have $U \subseteq [0,1]^3$, where the three coordinates correspond to the sets $\{1\}$, $\{2\}$ and $\{1,2\}$ (as we remarked above, the empty set plays no role). For a graphon $W$, we define the set $U_W = \{(x_1, x_2, x_{12}) \in [0,1]^3 : \; x_{12} \le W(x_1,x_2)\}$. Then it is easy to see that $t(F, U_W) = t(F,W)$ for any simple graph $F$.

Elek and Szegedy [2012] prove the following.

Theorem 23.12. *For every convergent sequence $(H_n)$ of $r$-graphs there is a hypergraphon $U$ such that $t(F,H_n) \to t(F,U)$ for every $r$-graph $F$.* □

The limit graphon is essentially unique up to some "structure preserving transformations", which are more difficult to define than in the case of graphs and we don't go into the details. Elek and Szegedy [2012] give several applications of Theorem 23.12. For a given hypergraphon $U$, they define $U$-random hypergraphs and prove that they converge to $U$. They derive from it the Hypergraph Removal Lemma due to Frankl and Rödl [2002], Gowers [2006], Ishigami [2006], Nagle, Rödl and Schacht [2006] and Tao [2006a]. As a refreshing exception, the statement of this lemma is a straightforward generalization of the Removal Lemma for graphs (Lemma 11.64); the proof of Elek and Szegedy is similar to our second proof in Section 11.8. They also derive the Hypergraph Regularity Lemma using Theorem 23.12, using a stepfunction approximation of hypergraphons.

This brings us to the Hypergraph Regularity Lemma, a very important but also quite complicated statement. There are several essentially equivalent, but not trivially equivalent forms, due to Frankl and Rödl [1992], Gowers [2006, 2007], Rödl and Skokan [2004], Rödl and Schacht [2007a, 2007b]. Proving the appropriate Counting Lemma for these versions is a further difficult issue, and I will not go into it. But I must not leave this topic without stating at least one form, based on the formulation of Elek and Szegedy [2012], which in fact generalizes the strong form of the Regularity Lemma (Lemma 9.5).

We have to define what we mean by "regularizing" a hypergraph. For $\varepsilon, \delta > 0$ and $k \in \mathbb{N}$, we define an *$(\alpha,\beta,k)$-regularization* of a $r$-graph $H$ on $[n]$ as follows. For every $i \in [r]$, we partition the complete hypergraph $K_n^i$ into $r$-graphs $G_{i,1}, \dots, G_{i,k}$. Let us think of the edges in $G_{i,j}$ as colored with color $j$. This defines a partition $\mathcal{P}$ of the edges of $K_n^r$, where two $r$-sets are in the same class if the colorings of their subsets are isomorphic. The family $\{G_{i,j} : \; i \in [r], j \in [k]\}$, together with an $r$-graph $G$ on $[n]$ will be called an $(\alpha,\beta,k)$-regularization of $H$, if

(a) every $r$-graph $G_{i,j}$ has quasirandomness at most $\alpha$, and

(b) $G$ is the union of some of the classes of $\mathcal{P}$, and

(c) $|E(H) \triangle E(G)| \le \beta\binom{n}{r}$.

Now we can state one version of the Hypergraph Regularity Lemma.

Lemma 23.13 **(Strong Hypergraph Regularity Lemma)**. *For every $r \ge 2$ and every sequence $\epsilon = (\varepsilon_0, \varepsilon_1, ...)$ of positive numbers there is a positive integer $k_\epsilon$ such that for every $r$-graph $H$ there is an integer $k \le k_\epsilon$ such that $H$ has an $(\varepsilon_k, \varepsilon_0, k)$-regularization.*

The main point is that to regularize $H$, we have to partition not only its node set, but also the set of $i$-tuples for all $i \le r$. Just like in the graph case, we could demand that the $i$-graphs $G_{i,j}$ have almost the same number of edges for every fixed $i$. Of course, the prize we have to pay for stating a relatively compact version is that it takes more work to apply it; but we don't go in that direction.

The extension of the theory exposed in this book to hypergraphs is not complete, and there is space for a lot of additional work. Just to mention a few loose ends, it seems that no good extension of the distance $\delta_\square$ has been found to hypergraphs (just as in the case of limit objects or the regularity lemma, the first natural guesses are not really useful). Another open question is to extend these results to nonuniform hypergraphs, with unbounded edge-size. The semidefiniteness conditions for homomorphism functions can be extended to hypergraphs (see e.g. Lovász and Schrijver [2008]), but perhaps this is just the first, "naive" extension. One area of applications of these conditions is extremal graph theory. The work of Razborov [2010] shows that generalizations of graph algebras and of the semidefiniteness conditions can be useful in extremal hypergraph theory. However, we have seen that graph algebras can be defined in the setting of gluing along nodes and also along edges, and this indicates that for hypergraphs a more general concept of graph algebras may be useful.

## 23.4. Categories

The categorial way of looking at mathematical structures is quite prevalent in many branches of mathematics. In graph theory, the use of categories (as a language and also as guide for asking question in a certain way) has been practiced mainly by the Prague school, and has lead to many valuable results; see e.g. the book by Hell and Nešetřil [2004].

One can go a step further and consider categories (with appropriate finiteness assumptions) as objects of combinatorial study on their own right. After all, categories are rather natural generalizations of posets, and there is a huge literature on the combinatorics of posets. However, surprisingly little has happened in the direction of a combinatorial theory of categories; some early work of Isbell [1991], Lovász [1972] and Pultr [1973], and the more recent work of Kimoto [2003a, 2003b] can be cited.

Working with graph homomorphisms, we have found not only that the categorial language suggests very good questions and a very fruitful way of looking at our problems, but also that several of the basic results about graph homomorphism and regularity can be extended to categories in a very natural way. The goal of this section is to describe these generalizations, and thereby encourage a combinatorial study of categories. (Appendix A.8 summarizes some background.)

**23.4.1. Cancellation laws.** Counting homomorphisms has been a main tool for proving cancellation laws for finite relational structures in Section 5.4, and it is not surprising that these results can be extended to locally finite categories (Lovász [1972], Pultr [1973]). The following two theorems generalize Theorem 5.34, Proposition 5.35(b) and Lemma 5.38 to categories.

THEOREM 23.14. *Let $a$ and $b$ be two objects in a locally finite category such that the direct powers $a^{\times k}$ and $b^{\times k}$ exist and are isomorphic. Then $a$ and $b$ are isomorphic.*

THEOREM 23.15. *Let $a, b, c$ be three objects in a locally finite category $\mathcal{K}$ such that the direct products $a \times c$ and $b \times c$ exist and are isomorphic.*

(a) *If both $a$ and $b$ have at least one morphism into $c$, then $a$ and $b$ are isomorphic.*

(b) *There exists an isomorphism from $a \times c$ to $b \times c$ that commutes with the projections of $a \times c$ and $b \times c$ to $c$.* □

So if there is any isomorphism $\sigma$ in Figure 23.3, then there is one for which the diagram commutes.

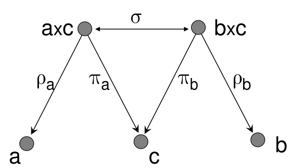



FIGURE 23.3

**23.4.2. Connection matrices and algebras of morphisms.** For the next theorem, we assume that $\mathcal{K}$ is a locally finite category that has a zero object, a left generator, pushouts, and epi-mono decompositions. Let $f$ be a real valued function defined on the objects, invariant under isomorphism. We say that $f$ is *multiplicative over coproducts*, if $f(a \oplus b) = f(a)f(b)$ for any two objects $a$ and $b$.

For every object $a$, we define a (possibly infinite) symmetric matrix $M(f, a)$, whose rows and columns are indexed by morphisms in $\mathcal{K}_a^{\text{in}}$, and whose entry in row $\alpha$ and column $\beta$ is $f\big(t(\alpha \vee \beta)\big)$ (since $\alpha \vee \beta$ is determined up to isomorphism, this is well defined). Note that specializing to the category of graph homomorphisms, $M(f, a)$ corresponds to the multiconnection matrix; to get the simple connection matrix, we have to restrict the row and column indices to monomorphisms.

One can extend the characterization of homomorphism functions in Corollary 5.58 to categories (Lovász and Schrijver [2010]); this theorem will also contain the dual characterization Theorem 5.59.

THEOREM 23.16. *Let $\mathcal{K}$ be a locally finite category that has a zero object $z$, a left generator, pushouts, and epi-mono decompositions. Let $f$ be a function defined on the objects, invariant under isomorphism. Then there is an object $b$ such that $f = |\mathcal{K}(., b)|$ if and only if the following conditions are fulfilled:*

(F1) $f(z) = 1$,

(F2) *$f$ is multiplicative over coproducts, and*

(F3) *$M(f, a)$ is positive semidefinite for every object $a$.*

We note that if there is an epimorphism from $a$ to $b$, then $M(f, b)$ is a submatrix of $M(f, a)$. Thus it would be enough to require the semidefiniteness condition for a left-cofinal subset of elements $a$.

COROLLARY 23.17. *Conditions* (F1)–(F3) *of the theorem imply that* (a) *the values of $f$ are non-negative integers,* (b) *the rank of $M(f, a)$ is finite for every $a$.*

Statement (a) of this corollary contrasts it with Theorem 5.54, where (thanks to the weights) the function values can be arbitrary real numbers. An analogue of (b) must be imposed as an additional condition e.g. in the characterization in Theorem 5.54, while in this version it follows from the other assumptions.

The conditions are very similar to those in Theorem 5.54, except that there the graphs cannot have loops and the matrices are indexed by monomorphisms only. As a consequence, the characterization concerns homomorphism numbers into weighted graphs, which has not been extended to categories so far.

The proof of Theorem 23.16 is built on similar ideas as the proof of Theorem 5.54 in Chapter 6, using algebras associated with the category. Since it is instructive how such algebras can be defined, we describe their construction below; for the details of the proof, we refer to the paper of Lovász and Schrijver [2010].

For two objects $a$ and $b$ in a locally finite category $\mathcal{K}$, a formal linear combination (with real coefficients) of morphisms in $\mathcal{K}(a,b)$ will be called a *quantum morphism*. Quantum morphisms between $a$ and $b$ form a finite dimensional linear space $\mathcal{Q}(a,b)$. Let

$$x = \sum_{\varphi\in\mathcal{K}(a,b)} x_\varphi \varphi \in \mathcal{Q}(a,b) \qquad \text{and} \qquad y = \sum_{\psi\in\mathcal{K}(b,c)} y_\psi \psi \in \mathcal{Q}(b,c),$$

then we define

$$xy = \sum_{\substack{\varphi\in\mathcal{K}(a,b)\\ \psi\in\mathcal{K}(b,c)}} x_\varphi y_\psi \varphi\psi \in \mathcal{Q}(a,c).$$

With this definition, quantum morphisms form a category $\mathcal{Q}$ on the same set of objects as $\mathcal{K}$. (Of course, $\mathcal{Q}$ is not locally finite any more, but it is locally finite dimensional.)

We can be more ambitious and take formal linear combinations of morphisms in $\mathcal{K}_a^{\text{out}}$ (for a fixed object $a$), to get a linear space $\mathcal{Q}_a^{\text{out}}$. This space will be infinite dimensional in general, but it has interesting finite dimensional factors. For each object $a$, the pushout operation $\wedge$ defines a semigroup on $\mathcal{K}_a^{\text{out}}$. Let $\mathcal{Q}_a^{\text{out}}$ denote its semigroup algebra of all formal finite linear combinations of morphisms in $\mathcal{K}_a^{\text{out}}$. So $\mathcal{Q}_a^{\text{out}} = \bigoplus_b \mathcal{Q}(a,b)$.

Just as in the case of graphs, every function $f:\ \mathrm{Ob}(\mathcal{K}) \to \mathbb{R}$ defines an inner product on $\mathcal{Q}_a^{\text{out}}$, by $\langle\alpha,\beta\rangle = f\big(h(\alpha\wedge\beta)\big)$. Condition (F3) in Theorem 23.16 implies that this inner product is positive semidefinite. Factoring out its kernel, we get a Frobenius algebra, which is finite dimensional (this takes a separate argument, since unlike in the proof of Theorem 5.54, this is not assumed directly). The proof of Theorem 23.16, just like the proof of Theorem 5.54, is built on studying the idempotent bases in these algebras.

EXAMPLE 23.18 **(Graph algebras)**. If the category is the category of graph homomorphisms, and $a$ is the $k$-labeled graph with $k$ nodes and no edges, then $\mathcal{Q}_a^{\text{out}}$ is the gluing algebra of $k$-multilabeled graphs. ◆

EXAMPLE 23.19 **(Flag algebras)**. Razborov's "flag algebras" [2007] can be defined in our setting as follows. We consider the category of embeddings (injective homomorphisms) between graphs. Fixing a graph $F$ (which Razborov calls a "type"), the morphisms from $F$ correspond to graphs with a specified subgraph isomorphic with $F$ (which Razborov calls a "flag"). The pushout of two such morphisms results in an object obtained by gluing together the two graphs along the image of $F$, which is exactly how Razborov defines the product in flag algebras. So flag algebras are the algebras $\mathcal{Q}_F^{\text{out}}$ in the category of monomorphisms between graphs. This is a subalgebra of the algebra $\mathcal{Q}_F^{\text{out}}$ defined in terms of all homomorphisms between graphs. ◆

**23.4.3. Regularity Lemma for categories.** There are more results on graph homomorphisms that extend quite naturally to the categorial setting. Let us state a generalization of the Regularity Lemma—both in its weak and original form (Lovász [Notes]).

To motivate the definitions below, consider a weighted graph $G$. This can be viewed as a weighting of the edges of a complete graph, i.e., as a quantum morphism $K_2 \to \tilde{K}_n$, which is symmetric, i.e., it is invariant under swapping the two nodes of $K_2$. Regularity lemmas try to find a partition (a morphism $K_n^\circ \to \tilde{K}_k$), and weighting of edges of $d = \tilde{K}_k$ (a quantum morphism in $\mathcal{Q}(a,d)$), such that "pulling back" these weights to $K_n^\circ$, we get a good approximation in the cut norm. How to translate to the categorial language that the cut norm of a weighted graph is small? It means that for every morphism $K_n^\circ \to K_2^\circ$, if we push forward the edgeweights, then the resulting edgeweights of $K_2^\circ$ are all small (this says that versions (a) and (c) of the cut norm in Exercise 8.4 are small, but these are all equivalent up to absolute constant factors). These considerations motivate the following general definitions.

Let $\alpha \in \mathcal{K}(a,b)$ and $\beta \in \mathcal{K}(c,b)$. We define a quantum morphism $\alpha\beta^* \in \mathcal{Q}(a,c)$ by

$$\alpha\beta^* = \sum_{\varphi \in \mathcal{K}(a,c):\, \varphi\beta = \alpha} \varphi.$$

This operation extends linearly to define $xy^*$ for $x \in \mathcal{Q}(a,b)$ and $y \in \mathcal{Q}(c,b)$. It is not hard to check that $x(zy)^* = (xy^*)z^*$, and $\langle x, yz^* \rangle = \langle xz, y \rangle$.

For every quantum morphism $x = \sum_\varphi x_\varphi \varphi \in \mathcal{Q}(a,b)$ and every object $c$, we define the *c-norm* of $x$ by

$$\|x\|_c = \max_{\beta \in \mathcal{K}(b,c)} \frac{\|x\beta\|_\infty}{|\mathcal{K}(a,b)|}.$$

This norm generalizes the cut norm: if $a = K_2$ and $c = K_2^\circ$, then a symmetric quantum morphism $x \in \mathcal{Q}(a,b)$ is a weighting of the edges of $b$, and it is not hard to see that $\|x\|_\square/2 \le \|x\|_c \le \|x\|_\square$.

Let $c^m$ denote the $m$-th direct power of the object $c$. The first inequality in the following lemma generalizes the Frieze–Kannan Weak Regularity Lemma 9.3, while the second implies the Original Regularity Lemma of Szemerédi 9.2.

LEMMA 23.20. *Let $\mathcal{K}$ be a locally finite category having finite direct products. Let $a$, $b$ and $c$ be three objects in $\mathcal{K}$, and let $m \ge 1$. Then for every $x \in \mathcal{Q}(a,b)$ there exists a morphism $\varphi \in \mathcal{K}(b,c^m)$ and a quantum morphism $y \in \mathcal{Q}(a,c^m)$ such that*

$$\|x - y\varphi^*\|_c \le \frac{1}{\sqrt{m}}\|x\|_2$$

*and*

$$\|x - y\varphi^*\|_{c^{2^m}} \le \frac{1}{\sqrt{\log^* m}}\|x\|_2.$$

The Weak Regularity Lemma is obtained, as described above, by taking $a = K_2$ and $c = K_2^\circ$ and applying the first bound. Note that a morphism in $\mathcal{K}(b,c^m)$ corresponds to a partition of $V(G)$ into $2^m$ classes. The Original Regularity Lemma can be derived from the second bound similarly. Strong versions can be generalized as well, but for the details we refer to Lovász [Notes].

There are many unsolved questions here: can the Counting Lemma be generalized to categories? Do the notions of convergence and limit objects be formulated in an interesting way? Could these results shed new light on hypergraph limits and regularity lemmas? Or perhaps even on sparse regularity lemmas?

Exercise 23.21. Let $\mathcal{K}$ be a locally finite category, and let $c$ be an object. Prove that every monomorphism in $\mathcal{K}(c, c)$ is an isomorphism.

Exercise 23.22. Let $\mathcal{K}$ be a locally finite category, and let $c, d$ be two objects. Suppose that there are monomorphisms in $\mathcal{K}(c, d)$ and in $\mathcal{K}(d, c)$. Prove that $c$ and $d$ are isomorphic.

Exercise 23.23. Let $\mathcal{K}$ be a locally finite category, and let $c$ and $d$ be two objects. For any two morphisms $\alpha \in \mathcal{K}(a, a')$ and $\beta \in \mathcal{K}(b, b')$, let $N_{\alpha,\beta}$ denote the number of 4-tuples of morphisms $(\varphi, \psi, \mu, \nu)$ $(\varphi \in \mathcal{K}(c, a), \psi \in \mathcal{K}(c, b), \mu \in \mathcal{K}(a', d), \nu \in \mathcal{K}(b', d))$ such that $\varphi\alpha\mu = \psi\beta\nu$. Prove that the matrix $N = (N_{\alpha,\beta})$, where $\alpha$ and $\beta$ range over all morphisms of the category, is positive semidefinite.

Exercise 23.24. Let $a$ and $b$ be two objects in a locally finite category. Suppose that the direct powers $a \times a$ and $b \times b$ exist and are isomorphic. Prove that $a$ and $b$ are isomorphic.

Exercise 23.25. Let $a, b, c, d$ be four objects in a locally finite category $\mathcal{K}$ such that the direct products $a \times c$, $b \times c$, $a \times d$ and $b \times d$ exist, $a \times c$ and $b \times c$ are isomorphic, and $d$ has at least one morphism into $c$. Prove that $a \times d$ and $b \times d$ are isomorphic.

## 23.5. And more...

There are many types of discrete structures for which one can try to define convergence and limit objects for growing sequences. This is typically not straightforward, as one can see from the case of simple graphs with (say) $\Theta(n^{3/2})$ edges. However, this approach has been successful in some cases.

It is a natural question to extend the theory of graph limits to directed graphs. Let us assume that these graphs are simple, so that there are no loops and there is at most one edge between two nodes in a given direction. Diaconis and Janson show that at least some of the theory can be developed based on the theory of exchangeable arrays (see Section 11.3.3). The limit object is a bit more complicated, it can be described by four measurable functions $W_{0,0}, W_{0,1}, W_{1,0}, W_{1,1} : [0, 1]^2 \to [0, 1]$ such that $W_{0,0}$ and $W_{1,1}$ are symmetric, $W_{0,1}(x, y) = W_{1,0}(y, x)$ and $W_{0,0} + W_{0,1} + W_{1,0} + W_{1,1} = 1$. The function $W_{0,1}(x, y)$ measures the density of edges from an infinitesimal neighborhood of $x$ to an infinitesimal neighborhood of $y$ etc. Some further remarks and observations can be found scattered in papers, but no comprehensive treatment seems to be known. Perhaps most of the extension is rather straightforward (but be warned: the theory of *existence* of homomorphisms between digraphs is much more involved—one can say richer—than for undirected graphs; see Hell and Nešetřil [2004]).

Posets can be considered as special digraphs, but they are sufficiently important in many contexts to warrant a separate treatment. Janson [2011a, 2012] starts a limit theory of posets. The treatment is based on methods similar to the limit theory of dense graphs in this book, but there are some analytic complications and interesting special features, for which we refer to the paper.

Going away from graphs, let us consider the set $S_n$ of permutations of the set $[n]$. Cooper [2004, 2006] defined and characterized quasirandomness for permutations, and proved a regularity lemma for them. Hoppen, Kohayakawa, Moreira, Ráth and Menezes Sampaio [2011, 2011] defined convergent sequences of permutations, and described their limit objects. Given a permutation $\pi \in S_n$ and a subset $A =$

$\{a_1, \dots, a_k\} \subseteq [n]$, we can define a permutation $\pi[A] \in S_k$ by letting $\pi[A]_i < \pi[A]_j$ iff $\pi_{a_i} < \pi_{a_j}$. For a permutation $\tau \in S_k$, let $\Lambda(\tau, \pi)$ denote the number of sets $A$ with $\pi[A] = \tau$, and define the density of $\tau$ in $\pi$ by $t(\tau, \pi) = \Lambda(\tau, \pi)/\binom{n}{k}$. A sequence of permutations $\pi_1, \pi_2, \dots$ (on larger and larger sets) is *convergent*, if for every permutation $\tau$, the number $t(\tau, \pi_n)$ tends to a limit as $n \to \infty$. Every convergent permutation sequence has a limit object in the form of a coupling measure on $[0, 1]^2$, which is uniquely determined. Král and Pikhurko [2012] have used this machinery of limit objects to prove a conjecture of Graham on permutations.

I have already mentioned the limit theory of metric spaces due to Gromov [1999]. While developed with quite different applications in mind, this turns out to be closely related to our theory of graph limits. Gromov considers metric spaces endowed with a probability measure, and defines distance, convergence and limit notions for them. A simple graph $G$ can be considered as a special case, where the distance of two adjacent nodes is $1/2$, the distance of two nonadjacent nodes is 1, and the probability distribution on the nodes is uniform. Under this correspondence, our notion of graph convergence is a special case of Gromov's "sample convergence" of metric spaces. Vershik [2002, 2004] considers random metric spaces on countable sets, and defines and proves their universality. He also characterizes isomorphism of metric spaces with measures in terms of sampling, analogously to Theorem 13.10. In a recent paper, Elek [2012b] explores this connection and shows how Gromov's notions imply results about graph convergence, and also how results about graph limits inspire answers to some questions about metric spaces. Perhaps Gromov's theory can be applied to graph sequences that are not dense, using the standard distance between nodes in the graph.

One of the earliest limit theories is John von Neumann's theory of continuous geometries. The idea here is that if we look at higher and higher dimensional vector spaces over (say) the real field, then the obvious notion of their limit is the Hilbert space. But, say, we are interested in the behavior of subspaces whose dimension is proportional to the dimension of the whole space. Going to the Hilbert space, this condition becomes meaningless. Neumann constructed a limit object, called a *continuous geometry*, in which the "dimensions" of subspaces are real numbers between 0 and 1. This construction can be extended to certain geometric lattices (Björner and Lovász [1987]), but its connection with the theory in this book has not been explored.

Perhaps most interesting from the point of view of quasirandomness and limits are sequences of integers, due to their role in number theory. (After all, Szemerédi's Regularity Lemma was inspired by his solution of the Erdős–Turán problem on arithmetic progressions in dense sequences of integers.) Often sequences are considered modulo $n$; this gives a finite group structure to work with, while one does not lose much in generality. Ever since the solution of the Erdős–Turán problem for 3-term arithmetic progressions by Roth [1952], through the general solution by Szemerédi [1975], through the work of Gowers [2001] on "Gowers norms", to the celebrated result of Green and Tao [2008] on arithmetic progressions of primes, a central issue has been to define and measure how random-like a set of integers is. I will not go into this large literature; Tao [2006c] and Kra [2005] give accessible accounts of it. What I want to point out is the exciting asymptotic theory of structures consisting of an abelian group together with a subset of its elements, and more generally, abelian groups with a function defined on them. There has been a lot of parallel developments in this area, most notably the work of Green, Tao and Ziegler [2011] and of Szegedy [2012a]. Not surprisingly, the latter is closer to the point of view taken in this book, and develops a theory of limit objects of functions on abelian groups, which is full of surprises but also with powerful results. (For example, to describe the limits of abelian groups, non-abelian groups are needed!) The theory has connections with number theory, ergodic theory, and higher-order Fourier analysis. This explains why I cannot go into the details, and can only refer to the papers.